The Complementarity of Human Capital and Language Capital in Foreign Direct Investment

Konara, P

10.1016/j.ibusrev.2018.10.009
International Business Review
Elsevier

The Complementarity of Human Capital and Language Capital in Foreign Direct Investment

Abstract

Integrating the literature on language-MNEs (multinational enterprises) in international business and economic theory of human capital (HC), we establish an analytical framework to systematically examine how HC and language capital (LC) jointly determine foreign direct investment (FDI). We contend that the extent to which MNEs can leverage HC in a host country for FDI depends on LC. Based on an extensive bilateral dataset covering 3,315 country pairs during 1995-2008, we reveal clear evidence on the moderating role played by LC in HC-FDI relationship and such evidence is robust to different measures used for different variables, the inclusion of more control variables and different samples.

Keywords

Language Capital; Human Capital; Foreign Direct Investment

The Complementarity of Human Capital and Language Capital in Foreign Direct Investment

1. Introduction

Multinational enterprises (MNEs) undertaking foreign direct investment (FDI)[1] need host country nationals (HCNs) and often have to deal with multiple languages. To this end, both the endowment of human capital (HC) and languages are expected to determine FDI and the degree of using languages at ease by MNEs could influence the effects of HC on FDI. Theoretically, there is general consensus of the positive impact of HC on FDI, but the empirical evidence is rather inconclusive. Positive impact is found, for instance, in Carstensen and Toubal (2004), Cleeve, Debrah and Yiheyis (2015), Gao (2005), Globerman and Shapiro (2002), Kinda (2010), Noorbakhsh, Paloni and Youssef (2001) and Suliman and Mollick (2009), while negative, little or no effect is revealed in Cheng and Kwan (2000), Kim and Park (2013), Nunnenkamp and Spatz (2002), Root and Ahmed (1979), Schneider and Frey (1985) and Wood, Mazouz, Yin and Cheah (2014). We argue that mixed findings could be the failure to consider the complementary role of language to HC in a cross-country setting. Though the importance of language is readily accepted, where and how it matters and what is the range of its effects have been "a surprisingly neglected subject of research in international business studies" (Piekkari, Welch & Welch, 2014, p. 9). To the best of our knowledge, there is little scholarship explicitly examining the role of HC and language in FDI in an integrated framework. This paper aims to fill the research gap by considering HC and language capital (LC), though two interrelated concepts, are conceptually distinct and can be analytically separated.

HC has long been discussed from a variety of perspectives in Economics, Management and Psychology and at various levels (individual-, firm- and country-level) (Ployhart & Moliterno, 2011).

Following Becker (1993) and Schultz (1961), we treat HC as a macro-, country-level concept and define it as productive investments embodied in human agents that improve knowledge and skills. In keeping with this definition, we operationalize HC in the form of education and training[2], "the most important investments in human capital" (Becker, 1993, p. 17), in a host country that can be used by MNEs as a productive resource for FDI. Education and training may result in a workforce that is literate, numerate, creative and skilled in the use and development of modern facilities and techniques. Such investments pay dividends for a country in its efforts to develop economic viability and improve investment climate to attract investors, foreign and domestic (McMahon, 2002).

The second building block of this study is language capital (LC). We extend the definition of LC at the firm level in recent literature (Welch & Welch, 2008, 2015)[3] to the country level and define it as a form of capital engendered from a country's language resources. It is capital because investing in language resources can lead to profit and can offer substantial transaction and consumption values. Like other forms of capital, e.g. financial capital and HC, LC is long-lived assets into which investments can be made for future returns. LC can be mobilized to facilitate actions and achieve a variety of positive outcomes. For the purpose of this paper, the focus is not on general language skills, but rather on language resources relevant at the interface between MNEs and the host country. It is, for example, the recognition of the benefits that a country can gain from investing in foreign languages education prompts countries, such as China, Japan and many European countries, to adopt English promotion policies in secondary and higher education. We consider LC as a separate construct to HC, albeit there is some degree of overlap between them. For example, having language skills in Spanish is part of HC, but it only represents LC for FDI originating from Spanish-speaking countries or countries that use languages close to Spanish according to language family affiliations.

Integrating the literature on language-MNEs in international business and the economic theory of HC, this paper contributes to the theoretical foundation of the research on HC-FDI relationship by theorizing the complementarity of HC and LC in determining FDI. Relying on the aggregation of human behavior theory to establish an analytical framework, we argue that the extent to which HC of HCNs can be utilized by MNEs depends on LC. Thus, this paper provides a step forward in understanding determinants of FDI by arguing for a more holistic consideration of HC and language in FDI location decision of MNEs and suggesting that conventional treatment of HC as a locational advantage should be augmented to incorporate LC. It is clear from recent survey and survey-type articles (Blonigen, 2005; Blonigen & Piger, 2014; Chakrabarti, 2001; Eicher, Helfman & Lenkoski, 2012), though HC and language have both been considered as determinants of FDI, their interplay has received no attention in the extant FDI literature. This research echoes an important note in Blonigen and Piger (2014) that "FDI decisions by MNEs are complex enough that interaction between key variables may be necessary to account for nonlinear effects of these variables on FDI patterns" (p. 776). By considering the complementary role of LC to HC in a cross-country setting, this research helps reconcile the mixed findings on the HC-FDI relationship.

In answering to recent call for sophisticated investigation of the role of language in international business (e.g. Brannen, Piekkari & Tietze, 2014; Harzing & Pudelko, 2013; Tenzer, Terjesen & Harzing, 2017), we position language at the center of this study and consider the LC concept in an inclusive manner. LC can be captured by language differences/similarities between the home and the host country. Most empirical studies considering language in FDI studies employ common language, thus focusing on MNE's home and host country speaking one and the same language. However, in multilingual countries, a number of languages are spoken, but by different proportions of population. Thus, there are different levels of *language overlap* across countries with common languages. In the absence of common language, lingua franca can be used or host country language(s) can be learnt.

Although all the variables discussed here can be used as proxies for LC, not all of them capture the same aspect, which means that information may be lost by including them separately in the regressions. Efficiency could be obtained by combining variables into composite indices[4]. We therefore use a variety of measures for LC, attempting to offer, to date, the most inclusive treatment of language in econometric studies of FDI. Our investigation should help depict the robust effects of LC on HC-FDI relationship and improve our understanding of the economic significance of language in global commerce (e.g. Adserà & Pytliková, 2015; Cook & Liu, 2016; Grin, 1994; Ku & Zussman, 2010; Melitz & Toubal, 2014).

In the next section, we first briefly review the literature on HC-FDI. This is followed by the introduction of LC concept and the formulation of hypothesis concerning the complementary relationship between HC and LC. The subsequent sub-section discusses the operationalization of LC. Section 3 explains data and methodology and section 4 presents results and robustness checks. The closing section offers policy and managerial implications of studying a phenomenon that is overlooked in the literature: that's, linking LC with HC and investigating their complementarity in MNE's FDI decision.

2. Literature Review

2.1 Human Capital (HC) and FDI

HC has been extensively studied in many disciplines (Ployhart & Moliterno, 2011). Following Becker (1993) and Schultz (1961), we define HC as productive investments embodied in human agents that improve knowledge and skills. The theoretical case for direct effects of HC on FDI has received much attention in the literature (Cleeve et al., 2015; Hughes, Powell, Chung & Mellahi,

2017; Noorbakhsh et al., 2001). Broadly speaking, the effects of HC can come about through knowledge and skills inherent in the HCNs that can be leveraged for FDI. HC helps MNEs build legitimacy for FDI in the host country, manage transaction costs in FDI, and transfer resources, particularly, knowledge to host country subsidiaries.

First, gaining external institutional legitimacy is essential for MNEs to succeed in a host country (Kostova & Zaheer, 1999). Legitimacy provides MNEs with the aptitude to overcome the liability of foreignness, realize acceptance and shield them against the unfavorable business climate. It is argued in the literature that MNEs employ HCNs to obtain legitimacy (Forstenlechner & Mellahi, 2011; Gong, 2003; Hughes et al., 2017). Hiring the "locals" enables MNEs to appear "local" and attuned with local institutions. HC is required by MNEs to develop a competence in dealing with host country institutional environment. HCNs can serve as bridges for MNEs to access local resources to address legitimacy requirements. MNEs can also gain internal legitimacy with ease when HCNs is better educated (Hughes et al., 2017). As Newburry, Gardberg and Belkin (2006) show, these HCNs are more likely to prefer working in MNEs, adjust to the MNEs operation/policies and work well with expatriates.

Second, because of bounded rationality, MNEs operating in a host country tend to face high transaction costs due to liability of foreignness, information costs, and the uncertainties of international transactions/operations (Rugman, 1986; Vidal-Suárez & López-Duarte, 2013). Transaction costs are closely associated with human agents and demand the input of HC to secure appropriate behaviors. To overcome costs associated with bounded rationality, MNEs require workforce with HC to secure appropriate information and to utilize information processing systems that determine the relevance of the acquired information (Rugman, 1986). Zaheer and Mosakowski (1997) consider that a major issue for MNEs is that they are "not being sufficiently embedded in the

information networks in the country of location". Information asymmetries increase transaction costs and create barriers for operations. Employing HCNs represents a channel through which MNEs can recognize and acquire relevant information to minimize transaction costs (Harzing, 2001; Vance & Paik, 2005; Vance, Vaiman & Andersen, 2009). An increase in the skill level of HCNs can also improve expatriate adjustment and performance which contributes to a reduction in transaction costs (Vance & Paik, 2005). Finally, establishing business practices to combat opportunism requires HC for contract design, modification and implementation. From the management perspective, Dow and Karunaratna (2006) note differences in education between countries "increase the risk and uncertainty of a manager properly understanding and communicating with the market, …, this is likely to increase the transaction costs and thus impact FDI market preferences" (p. 582).

Third, one advantage of MNEs over local firms is internal transfer of firm-specific assets which are often of intangible nature (Dunning & Lundan, 2008). Because of asset specificity and intangibility, MNEs need HC endowment of the host country to effectively exploit these firm-specific assets within their subsidiaries (Hitt, Biermant, Shimizu & Kochhar, 2001; Hughes et al., 2017). HC ensures the absorption and assimilation of resources/knowledge and their modification and further development in accordance with local conditions to commercial ends (Colakoglu, Yamao & Lepak, 2014; Reiche, Harzing & Pudelko, 2015; Uhlenbruck, 2004). Educated HCNs can learn and adapt firm-specific assets faster and more readily, providing time and cost advantages to the investing firms (Schomaker & Zaheer, 2014). Without sufficient level of HC endowment, unskilled workforce lacks the capability to utilize resources transferred to the subsidiaries (Contractor & Mudambi, 2008). When subsidiaries have inadequate "absorptive capacity", MNEs face extra costs on training and developing HC (Wood et al., 2014) and would also rely more on parent country nationals (Gong, 2003; Harzing, 2001), all of which can add to the transaction costs of FDI.

Finally, HC also indirectly affects FDI through increasing the attractiveness of local investment climate (Kinda, 2010). The extant economic development and economic growth literature has established macro-level effect of HC and education on improving social development, economic growth and productivity, which are important for attracting FDI (e.g. Becker, 1993; McMahon, 2002). HC is also an important input for generating new products or ideas that underlie technological progress and absorbing the new products or ideas that have been discovered elsewhere. Thus, HC contributes to the formation of a favorable and quality investment climate that is conducive to investment in general and FDI in particular. We therefore posit:

Hypothesis 1: Human capital has a positive impact on FDI.

Despite the general consensus on the theoretically expected positive relationship between HC and FDI, empirical evidence is inconclusive as explained in section 1. Such may indicate that the impact of HC needs to work through language in determining FDI in a cross-country setting. This is because much of the HC in a country is gained through the use of local languages. The economic value of HC, therefore, is linguistically bound. Assuming MNEs can utilize HC of HCNs at the full potential without taking into account language overstates the effects of HC.

2.2 The Complementarity of Language Capital and Human Capital in FDI

We define LC at the country level as a form of capital engendered from a country's language resources. LC shapes economic actions and outcomes by affecting the notion of shared values and identity (Harzing & Feely, 2008; Tenzer, Pudelko & Harzing, 2014) and by providing a medium through which economic agents can effectively communicate, interact and coordinate with each

other. Given the focus of this paper, LC is not about general language skills, but language resources relevant at the interface between MNEs and the host country.

LC is important in many FDI processes, from collecting information and undertaking business negotiations in preparation for FDI to transferring tacit knowledge to host country subsidiaries, continuously supervising, developing and monitoring different functions of these subsidiaries, managing local workforce and interacting with local suppliers and local customers after FDI is commenced. Language barriers increase the liability of foreignness (Berry, Guillén & Zhou, 2010; Luo & Shenkar, 2006), information costs and the uncertainties of international transactions (Dow & Karunaratna, 2006), and trigger problems of miscommunication, mistrust, frustration, conflict, and resistance among co-workers in MNEs and between the MNE personnel and those of the suppliers and customers (Barner-Rasmussen & Björkman, 2005; Harzing & Feely, 2008; Schomaker & Zaheer, 2014; Tenzer et al., 2014; Vidal-Suárez & López-Duarte, 2013).

In a cross-country setting, how HC can be effectively leveraged depends on the ability of MNE workforce to utilize LC to communicate, integrate socially, create/share tangible and intangible resources and develop/maintain trust among themselves and with external economic agents. A high level of LC offers MNEs linguistic ease and enables them to leverage HC in acquiring, presenting and interpreting information in order to distinguish between good and bad investments, improving post-investment management, and possibly increasing subsequent investments. The complementarity between HC and LC helps MNEs in legitimacy building, transaction cost reduction, and knowledge transfer.

First, as we established above, MNEs require HC to gain legitimacy, but the extent to which HC's role is effective depends on LC. MNE subsidiaries simultaneously come under external institutional

pressures from the host country environment and internal institutional pressures from within their organizations, this is termed as "institutional duality" by Kostova and Zaheer (1999). MNEs need LC to effectively utilize HC embedded in HCNs to understand and appreciate the host country's institutions and to design strategies aligning practices with both external and internal environment for gaining legitimacy. HC has been associated with more effective strategic choices and effective strategy implementation (Hughes et al., 2017). But in a multi-country setting, linguistic difficulties can undermine the value proposition of HC to MNEs by constraining the performance of employees or by obscuring their (both the actual and the perceived) professional competence (Aichhorn & Puck, 2017b; Tenzer et al., 2014). In contrast, if MNEs can use and/or acquire host country language with ease, they can effectively leverage HC as MNE employees can undertake appropriate/meaningful communication in their working relationships in intra-organizational and inter-organizational, and within country and between-country settings (Piekkari et al., 2014). With LC, MNEs can more effectively utilize HC to curb the host country's "continuing suspicion" towards them, at the same time, to offer the subsidiary legitimacy in the eyes of the parents (Kostova & Zaheer, 1999).

Second, like other firms, MNEs can leverage HC to deal with transaction costs. However, different from a single country context, to harness the value of HC, language plays an important role for MNEs. The combination of HC and LC would help manage opportunistic behaviors and reduce uncertainties and linguistic difficulties can exacerbate transaction costs related to bounded rationality and information asymmetry. Employees' ability to process complete information for actions and managers for decisions depend on their abilities to utilize language for communicating, socially interacting and building trust with other home and host country personnel. Language barriers would mean their decision-making processes, consequently the value of HC, being compromised. In order to overcome language barriers and maximize the potential of HC, MNEs can use translators, undertake language training and development or rely on a few multi-lingual employees who act as

language nodes between speakers of different languages (Luo & Shenkar, 2006). However, in any case, more costs are associated with FDI. In addition, using a translator may not always be effective as there is a risk of miscommunication because a translator may not have a level of proficiency in the specialist subject. The reliance on multi-lingual employees places an extra workload on them, which could compromise their performance in their formal functions (Feely & Harzing, 2003). In summary, when making FDI location decision, MNEs pay attention to transaction costs in a host country and LC strengthens the effect of HC on MNE's ability in dealing with transaction costs in FDI, therefore affecting FDI location decision.

Third, HC plays an important role in MNCs knowledge transfer. However, simply having HC is not sufficient. HC and LC need to work jointly in the processes of transferring resources, particularly knowledge, within MNE networks. Effective transfer requires clear communication through language (Peltokorpi, 2017; Reiche et al., 2015; Schomaker & Zaheer, 2014; Welch & Welch, 2015). In the context of institutional duality, knowledge transfer is also not straightforward. Though the transfer of knowledge from parents gives subsidiaries advantages over their competitors and the use of such knowledge grants them internal legitimacy in the eyes of the MNE parents, in view of external legitimacy, adaption is often required to align with the external institutional context because imposition of transferred knowledge in a different institutional setting can be potentially damaging (Hughes et al., 2017). LC determines the degree to which MNEs can fully utilize HC embedded in HCNs to accurately and efficiently engage in knowledge transfer. Thus, we can formulate the following hypothesis:

Hypothesis 2: Language capital positively moderates the positive impact of human capital on FDI.

2.3 Operationalizing Language Capital

To operationalize LC in the context of FDI study, we first consider the most obvious scenario, i.e. the host country speaks one and the same language as the home country of the MNEs. The shared language offers vehicle for "direct communication" between host and home country managers, employees and other economic agents (Melitz, 2008). A common language also reflects similarities in how people interpret, understand and respond to information provided by others, translate and construct ideas and issues, and facilitate the formation of identity and increases the frequency of communication (Harzing & Feely, 2008; Tenzer et al., 2014). Thus, the host country sharing common language as the home country has an advantageous position in LC.

Such common language indicator, however, only considers whether the language can be either officially or widely spoken in both countries. It does not reflect the fact that in multilingual countries, languages are used by different proportions of population. Thus, there are different levels of *language overlap* across countries with common languages. For example, English is the common language for both Canada and Switzerland. Relatively speaking, a higher percentage of the population in Canada speak English than that in Switzerland (Dow & Karunaratna, 2006). Thus, an UK MNE conducting FDI tends to face a higher degree of linguistic ease in Canada than in Switzerland as they can "directly communicate" with a higher proportion of English speakers in Canada than in Switzerland. We therefore expect LC of a host country is positively linked to language overlap between the home and host country of FDI.

In the case when the home and host country of MNEs do not share a common language, MNEs can achieve linguistic ease in a host country through using lingua franca, particularly English, and/or learning its language(s). Lingua franca, a non-native language in both the home and host country of

MNEs can be viewed as a vehicle language (Ku & Zussman, 2010), offering MNEs the opportunity to overcome language barriers in the host country. For example, a German MNE undertaking FDI in China or Japan can choose to employ a lingua franca. Among all languages, English has a dominant position not only in terms of its general use, but also its popularity in international business (Brannen et al., 2014; Harzing & Pudelko, 2013; Hejazi & Ma, 2011; Ku & Zussman, 2010). Moreover, a bulk of human knowledge is stored in English. As noted by Nunan (2003), "English is currently the undisputed language of science and technology, and scientific journals in many countries are now switching from the vernacular to English" (p. 590). In international business, English is not only used as corporate language by MNEs from English-speaking countries but also by MNEs from non-English-speaking countries. Studying more than 800 subsidiaries located in 13 host countries with headquarters (HQs) in more than 25 countries, Harzing and Pudelko (2013) find that 99.3% Anglophone MNEs, 93.2% of Nordic MNEs and 84.2% of Continental European MNEs use English in communication with HQs. Even among Asian MNEs, a sizable proportion has also used English (38.1%). Similar findings is shown by Barner-Rasmussen and Aarnio (2011). Indeed, MNEs from non-English speaking countries increasingly see employing English as an official language would help increase their global competitiveness and attract and retain global talent, despite the issues associated with language standardization (Harzing & Pudelko, 2013; Reiche et al., 2015). Given the prevalent use of English in MNEs, *whether or not a host country uses English* is closely linked to LC of the host country.

LC can also be associated with how easy or difficult a host country language can be learnt by MNE employees who overcome linguistic barriers through language acquisition. Linguists, considering structural differences across languages, argue that some languages are more similar than others in terms of language family affiliations[5], i.e. language pairs have different degree of linguistic proximity (Lewis, Simons & Fennig, 2014). Learning a host country language that is linguistically

distant from the mother tongue is harder and more costly than learning one that is linguistically proximate to the mother tongue (Adserà & Pytliková, 2015; Chiswick & Miller, 2005). Thus, we anticipate LC of a host country is positively linked to linguistic proximity.

Although all four variables mentioned above are proxies for LC, not all of them capture the same aspect. MNEs may benefit from multiple aspects of LC simultaneously. For example, an MNE investing in a host country that has English as official language as well as shares a common non-English language with its home-country may benefit more than another MNE in a host country that only has English as official language and no shared common non-English language. Multiple aspects of LC can be particularly useful when MNEs use more than one language for corporate communication (Aichhorn & Puck, 2017a; Peltokorpi, 2015). Existing evidence also show even when MNEs adopt one language for corporate communication, employees frequently switch between languages (Aichhorn & Puck, 2017a). Therefore, considering multiple aspects of language bridges can be useful, which requests the consideration of a composite index to capture LC.

3. Data and Research Methods

3.1 Sample and Variables

Bilateral FDI data were taken from OECD's International Direct Investment Database where FDI inflows to OECD countries and FDI outflows from OECD countries are reported but not bilateral flows between non-OECD countries. Nevertheless, the final sample still covers FDI for a large set of host and home countries, i.e. 3,315 country pairs representing 81 host countries and 88 home countries (Appendix I). The time-period covered by the sample is between 1995-2008. Given the global financial crisis in 2008-2009, FDI flows after 2008 are subject to the influence of this event.

To include data after 2008 will, therefore, require the shift of the focus of the paper to include the impact of financial crises, the literature of which "is still in its infancy and scattered" (Weitzel, Kling & Gerritsen, 2014, p. 213). Therefore, we consider that confining the sample to up to and include 2008 is appropriate.

Variables measurements and data sources are summarized in Appendix II. The dependent variable is bilateral FDI flow. We prefer FDI flows over FDI stocks[6]. The main reason is to do with the ways FDI stocks are calculated can vary across countries (Medvedev, 2012) and depend on financial markets and exchange rate volatility of host countries (Beugelsdijk, Hennart, Slangen & Smeets, 2010). FDI inflows are less vulnerable to such "book-value" bias inherited in FDI stocks (Noorbakhsh et al., 2001; Root & Ahmed, 1979). Additionally many of our independent variables are changing yearly, FDI stocks would not necessarily reflect the effect of these annual changes on FDI because FDI stocks can be accumulated over a long period of time (Medvedev, 2012).

3.1.1 Human Capital

We employ three variables of HC. The first is *Secondary education* measured by secondary school enrolment rate. The Economics perspective of HC at the macro level often links HC to education as we explained in Section 1. *Secondary education* is the most popular one used in the literature. This is partially because of data availability. But there is also a theoretical reasoning behind its popularity. Meier (1995) argues "the most critical manpower requirement tends to be for people with a secondary education who can be managers, administrators, professional technicians, or sub-professional technical personnel" (p. 315).

MNEs undertaking FDI also need HCNs with higher level qualifications given MNEs' characteristics of high levels of R&D, high level of utilization of professional and technical workers, new and complex products and advertising. From this perspective, our second HC variable is *Tertiary education* measured by tertiary education enrolment rate.

Finally, though education is often regarded as the most important element, HC can be associated with on-the-job training and health care (Becker, 1993). World Economic Forum published '*Higher education and training*' measures in the Global Competitiveness Index, a composite index based on education, quality of education and on-the-job-training, and we use this as our third HC variable. Arguably this variable may better reflect a country's HC stock than *Secondary education* and *Tertiary education*. However, the data only cover 2006-2008. Using this variable reduces sample size substantially. In what follows, we will first present results with *Secondary education* and carry out robustness checks using *Tertiary education* and *Higher education and training*.

3.1.2 Language Capital

The selection of key variables to represent LC is a major issue. We first adopt existing measures for variables discussed in section 2.3. *Common language* is a binary variable that reflects whether the host and home country share a common language.

Language overlap captures the probability that a randomly chosen individual from the host county will speak the language(s) of the home country. It is measured by the incidence of home country's dominant language(s) in the host country (i.e. the proportion of population (p) in the host country who are able to speak the major language(s) of the home country). We reverse coded Dow and Karunaratna (2006)'s 5-point language distance indicator, so that higher value reflects higher level of

linguistic ease. The five-point scale is coded as follows: 1 if p is less than 1%, 2 if p is greater than or equal to 1% but less than 5%, 3 if p is greater than or equal to 5% but less than 50%, 4 if p is greater than or equal to 50% but less than 90%, 5 if p is greater than or equal to 90%. For example, Canada has two major languages (English and French), UK has one major language (English) and France has one major language (French). Since the percentage of English and French speakers in Canada are approximately 61% and 23%, respectively (Lewis et al., 2013), values of the indicator variable considered are 4 for FDI flows from UK to Canada and 3 for FDI flows from France to Canada.

Linguistic proximity captures linguistic distance between the two closest major languages for each pair of countries. We adopt a reverse coded measure used in Dow and Karunaratna (2006). Linguistic proximity between the two closest major languages for each pair of countries is categorized to a five-point scale with lower value indicating a lower level of linguistic proximity (or a higher level of linguistic distance) between two languages, hence more difficulty in learning the host country language. The five-point scale is coded as follows: 1 if languages are form different families, 2 if languages are from the same family but different branches, 3 if languages are from the same branch but different at the first sub-branch level, 4 if languages are from the same sub-branch at the first level but different at the second level, and 5 if languages are same. For example, English and French belong to the Indo-European language family and Mandarin along with other languages spoken in China (e.g. Cantonese) belongs to Sino-Tibetan language family, therefore, values of the linguistic proximity indicator variable considered in this study are 2 for FDI flows between UK and France and 1 for FDI flows between France and China, reflecting the fact that, for a native French Speaker, learning Mandarin would be harder than learning English.

English as lingua franca is measured in two ways: whether English is recognized as an official language and whether English is used as a medium of instruction in the host country. The Central

Intelligence Agency (CIA)'s World Factbook provides the official language(s) for each country. We used this information to construct *English as official language* variable. English may not be treated as an official language, but may be incorporated into the education system. Many countries have recognized its role in globalization and have undertaken policy changes to introduce English into their education systems (Nunan, 2003). Increasingly, more countries are starting to use English as the medium of instruction in primary, secondary and tertiary education (Dearden, 2014). Universities in the US, the UK and Canada provide a list of countries that use English as the official language and language of instruction, and students coming from these countries are often exempt from the English Language Proficiency Requirements. Although this list may vary depending on the University profile and the origin, it can be useful in terms of identifying whether a particular country uses English as a medium of instruction. We use the exempted country list from the University of Florida to create *English as medium of instruction* variable.

As not all of the above variables capture the same aspect of LC, this means that information may be lost by including them separately in the regressions. Efficiency could be obtained by combining them. We consider two composite indices. First, we carried out factor analysis on all five measures to derive LC1 which is the first factor. This factor accounts for 80% of the variance of the five measures. Second, we use a categorical variable (LC2) to capture both *Common language* and English variables, taking zero if there is no common language including English between home and host country, one if only one of *Common language*, *English as official language* and *English as medium of instruction* variables is of value one and two if *Common language* is of value one and any one of *English* variables is of value one.

3.1.3 Control Variables

To ensure the observed relationship between HC, LC and FDI is not spurious, we include a wide range of FDI determinants following previous work. First, we control for other societal features using *Cultural distance*, *Religious distance* and *Historical links*. *Cultural distance* is a composite variable based on Hofstede's cultural dimensions and calculated using Mahalanobis method. This method accounts for variance of each dimension and co-variance between dimensions and produces a scale-invariant distance measure (Berry et al., 2010). Religion can be considered as a distinctive cultural feature (Gomez-Mejia & Palich, 1997). *Religious distance* approximates the religion aspect of cultural effects (Dow & Karunaratna, 2006). *Historical links*, colony relations in particular, promotes the spread of language, culture and religions (Makino & Tsang, 2011).

Guided by previous literature (Blonigen, 2005; Blonigen & Piger, 2014; Chakrabarti, 2001; Eicher et al., 2012), we further include: *population* and *GDP per capita* of the host country (Chakrabarti, 2001; Schneider & Frey, 1985), *home country GDP* (Chakrabarti, 2001), *trade openness* of the host country (Noorbakhsh et al., 2001; Suliman & Mollick, 2009), *infrastructure* in the host country (Suliman & Mollick, 2009), *geographical distance* and adjacency (*Border*) between the host and home country (Oh, Selmier & Lien, 2011), and whether the host and home country share a regional trade agreement (*RTA*) (Oh et al., 2011).

We also include two variables that have not received much focused attention in FDI literature. First, there is a vast literature on the consequences of legal origin, arguing that the legal origin of a country can determine its economic outcomes (La Porta, Lopez-de-Silanes & Shleifer, 2008; La Porta, Lopez-de-Silanes, Shleifer & Vishny, 2000). From an investment perspective, countries with common law systems (originating in English law), in comparison to those with civil law systems, tend to place greater emphasis on private property rights, are more protective of outside investors and are less associated with a heavy hand of government ownership and regulation, and hence may be

more attractive to investors (La Porta et al., 2008; La Porta et al., 2000). Therefore, following this line of reasoning, we can expect a link between legal origins and FDI location decision. We include *Legal origin* that identifies whether a country has a common-law system or a civil law system. Second, it is widely accepted that FDI is motivated by market liberalization (Meyer, Estrin, Bhaumik & Peng, 2009). However, the extant literature tends to focus on trade liberalization or trade openness, rather than broad institutional environment. Following Meyer et al. (2009), we examine institutions that support market efficiency and choose the following categories from Heritage Foundation Economic Freedom Index: (1) business freedom, (2) trade freedom, (3) property rights, (4) investment freedom, and (5) financial freedom. We follow Heritage's method and construct a country's market liberalization index (*Market liberalization*) by taking equal weighted average of the five categories.

3.2 Model

In econometric analysis, the first issue we need to address is related to non-positive values in FDI statistics. In our sample, 51% of the FDI observations are non-positive. Log transformation leaves out these values, potentially resulting in selection bias. That's, if zero FDI flows tend to occur between language distant countries, dismissing them would undermine the whole purpose of our research. Existing literature offers a couple of ways to deal with this issue: adding constant values to transform non-positive to positive values or taking the logged absolute values of the variable, incrementing it by one, then multiplying by the sign of the original variable (Dhanaraj & Beamish, 2009). However, given the large number of non-positive values in our sample and the range of FDI spanning from -48 billion US$ to 114 billion US$, both forms of transformations could introduce bias. We therefore opt to a linear model.

$$FDI_{ijt} = \beta_0 + \beta_1 LC_{ij} + \beta_2 HC_{it} + \beta_3 LC_{ij} x HC_{it} + \beta_4 Population_{it} + \beta_5 GDP\ per\ capita_{it} + \beta_6 GDP\ Home_{jt} + \beta_7 Trade\ openness_{it} + \beta_8 Infrastructure_{it} + \beta_9\ Geographical\ distance_{ij} + \beta_{10} Border_{ij} + \beta_{11} RTA_{ijt} + \beta_{12} Cultural\ distance_{ij} + \beta_{13} Religious\ distance_{ij} + \beta_{14} Historical\ links_{ij} + \beta_{15} Market\ liberalization_{it} + \beta_{16}\ Legal\ origin_i + Region_i + \mu_i + \mu_j + \delta_t + \gamma_{ij} + \varepsilon_{ijt}$$

where i, j and t indicates host country, home country and year, respectively. FDI_{ijt} is the bilateral FDI flows between host country i and home country j at time t. LC_{ij} x HC_{it} is the interaction term between language capital and human capital and a positive coefficient indicates support to our central hypothesis. $Region_i$ is a set of regional dummies to capture any region-specific issues that can affect FDI inflows. μ_i, μ_j and δ_t are host-country-specific and home-country-specific and year-specific effects respectively[7]. γ_{ij} represents country-pair specific (dyadic) effects. ε_{ijt} is the white noise disturbance terms. As guided by the literature and past empirical studies, we expect β_9, β_{12} and β_{13} to be negative, β_{11} to be indecisive, and coefficients of rest of the control variables to be positive in priori.

A number of other issues are further considered. First, in panel data estimation, μ_i, μ_j, δ_t and γ_{ij} can be treated as random variables (error component approach) or fixed parameters (fixed effects approach), depending on whether unobserved characteristics are correlated with explanatory variables. Given our interest on the time-invariant bilateral variable – LC, we cannot use dyadic-specific (γ_{ij}) fixed effects, but we can use fixed effects for home-country effects (μ_j), host-country effects (μ_i) and time-specific effects (δ_t) to account for host and home country heterogeneity and business cycle effects (Anderson & van Wincoop, 2003), a popular approach among existing studies. Since panels with long time series could have the problem of serial correlation, the Wooldridge test for autocorrelation was used and the results indicate no serial correlation. All estimations were calculated with robust standard errors to control for heteroscedasticity.

4. Results

TABLES 1 & 2 HERE

Table 1 presents descriptive statistics and correlation coefficients. It is worth noting that correlation coefficients between HC and LC variables are low, further supporting our treatment of the two key constructs separately.

Among LC variables, with the exception of correlation coefficient between *Common language* and *Language overlap* (0.58), that between *Language overlap* and *Linguistic proximity* (0.56) and that between *English as official language* and *English as medium of instruction* (0.96), all other coefficients are low. This therefore confirms that we should employ a variety of measures to capture LC.

There is high correlation among three HC measures (0.66-0.81), particularly between *Secondary education* and *Higher education and training* (0.81). Therefore, though *Higher education and training* may ideally appear to be a better proxy for HC than *Secondary education*, the latter can be used to approximate the former. As *Secondary education* has the most comprehensive data coverage, allowing us to utilize the full range of data, *Secondary education* is used for baseline regressions. The results are presented in table 2.

To investigate the moderating effects of LC on HC-FDI relationship, we focus on the interaction term (*LC*x*Secondary education*), with a positive coefficient indicating LC positively moderate the impact of HC on FDI. (2.1) and (2.2) use *Common language* and *Language overlap*, respectively. (2.3) and (2.4) include *English as official language* and *English as medium of instruction* to capture

English as lingua franca effects and the sample only cover country pairs that do not have a common language and non-English speaking home countries. As English facilitates cross-border activities both between English-speaking countries (the intra-language effect of English) and between countries that have different languages (the inter-language effect or lingua franca effect) (Hejazi & Ma, 2011), we also estimated this model for the full sample and the results are presented in (2.5) and (2.6). Comparing (2.3)-(2.6), we can clearly observe the importance of English for both native and non-native English-speaking countries. Finally, (2.7) shows the results using *Linguistic proximity*.

Across all specifications in table 2, *Secondary education* is either positive but statistically insignificant or negative. On the other hand, *LC*x*Secondary education* is consistently positive and significant, thus H1 is not supported but H2 is. Put it differently, the positive effects of HC on FDI must work through LC. We should also note that the coefficients on LC variables are negative and significant in most of the estimations, which indicates that there is a threshold effect: that is the HC variable as captured by secondary education would need to reach a certain threshold level before the interaction between HC and LC would be large enough to compensate the direct LC effects. To illustrate, for specification 2.3 in table 2, this threshold is equal to 53%. This means that, to capitalize on the effects of English as lingua franca for attracting FDI, secondary education enrolment rate needs to reach at least 53%.

Ceteris paribus, the marginal effect of HC under various conditions of different LC measures can be comprehended by examining the conditional slope of *Secondary education* at the values of various LC indicators[8]. The calculated results are reported in table 3. First, in the case of *Common language*, the marginal effect of *Secondary education* is higher by 4.7 when the host and home country do share a common language than when host and home country do not share a common language. Second, when LC effects are captured by *Language overlap*, the impact of *Secondary education* on

FDI increases from 2.5 to 12.7 when *Language overlap* moves from the lowest level (1) to the highest level (5). This means, when *Language overlap* is at its lowest value of one (i.e. less than 1% of the host country speak the languages of the home country), one percentage point increase in secondary education enrollment rate will increase FDI inflows by US$2.5 million, and when *Language overlap* is at its highest value of 5 (i.e. more than 90% of the host country speak the languages of the home country), one percentage increase in secondary education enrollment rate will increase FDI inflows by US$12.7 million. These increases are considerably large given that the average FDI inflow in our sample is US$310 million. Similarly, the impact of *Secondary education* increases from 1.4 to 7.3 within the range of *Linguistic proximity*. Considering English as lingua franca, the marginal effect increases by 8.5 as *English as official language* moves from zero to one, and that increases by 8.2 as *English as medium of instruction* moves from zero to one. In sum, across all LC measures, marginal effect of *Secondary education* increases considerably when LC of the host country increases.

TABLE 3 HERE

In terms of control variables in table 2, *Population*, *GDP per capita* and *GDP Home* have the expected positive sign and are significant in most of the estimations. *Geographical distance* has the expected negative sign and is highly significant. In terms of other control variables, *Border* has a positive coefficient and *Cultural distance* has a negative coefficient and they are significant in all specifications. *RTA* has a negative coefficient and significant, indicating that FDI and trade are largely substitutes rather than complimentary. *Legal origin* is positive and significant in some specifications. Other control variables are insignificant. Therefore, control variables have performed modestly well.

Although the Wald χ^2 is significant in all specifications suggesting a good fit to the data, R^2s indicate that the overall explanatory power of the models is modest. They are lower compared to existing studies on language and FDI, e.g. Hejazi and Ma (2011), Oh et al. (2011) and Selmier and Oh (2013). This could be because these studies have only considered positive FDI flows and take logs to smooth out the dependent variables. As we mentioned earlier, 51% of the FDI observations are non-positive and excluding them will bias the estimated results. To check this, we re-estimated our models for positive FDI flows (i.e. without considering non-positive FDI flows), and R^2s increase considerably.[9]

Next, we consider using *Higher education and training* and *Tertiary education* to measure HC, the estimation results are provided in tables 4 and 5, respectively. The coefficients on *LC*x*Higher education and training* in table 4 are again positive and statistically significant in all specifications. In table 5, the coefficients on *LC*x*Tertiary education* remain positive for all specifications and statistical significance also hold for five out of seven specifications. These results are in line with those in table 2, strongly reinforcing the message that LC and HC are complementary in determining FDI.

TABLES 4, 5 AND 6 HERE

So far, we have included different LC measures separately in regressions. To assess the combined effect of the different aspects of the LC measures, we repeated the estimations in Table 2 with two composite indices and the results are presented in table 6. The interaction term is again positive and statistically significant. We repeat the same exercise using two other measures of human capital and obtain similar results. Thus, MNEs can benefit from multiple aspects of LC simultaneously.

To further test robustness[10], first, we add more distance measures as control variables to baseline models. Five measures including demographic distance, economic distance, financial distance, knowledge distance, and political distance are taken from Berry et al. (2010). Since only a fraction of bilateral pairs from our total sample have all five distance measures, the sample size is reduced significantly. Second, we remove FDI inflows to the US and the UK from the sample. This is to make sure that our previous results are not driven by strong US and city of London effects as both the US and the UK are English-speaking countries. Third, we use factor analysis to derive a composite index for HC. Again, the results on key variables and the interaction term between HC and LC largely remain intact. In summary, empirical results strongly support the complementary role played by HC and LC in FDI regardless of the measures used for different variables, the inclusion of more control variables and the employment of different samples.

5. Conclusions

This paper attempts to probe the complementary role of HC and LC in determining FDI. Motivated by the current state of the literature, that's, the role of language in the extant FDI research is under-researched and the evidence on the relationship between HC and FDI is inconclusive, we build on the existing literature on language-MNEs and extend it in a new direction by theorizing the complementarity of HC and LC. The extent to which MNEs can leverage HC in host countries for FDI operations depends on LC. To capture LC effects, we consider not only host country languages that can be directly used by MNEs but also ones that can be acquired by MNEs. Using measures that represent major languages and total population of the countries and based on an extensive bilateral dataset covering 28 OECD countries and 88 partner countries during 1995-2008, we reveal strong evidence on the complementarity of HC and LC. Thus, the results clearly indicate HC is a necessary

but insufficient condition for FDI location decision. In a global context, the effects of HC are contingent on LC.

The findings have important policy relevance. FDI has been growing rapidly worldwide and is increasingly recognized to play a significant role in economic development. National policymakers seeking to attract inward FDI or promote outward FDI should pay attention not only on education and the development of local workforce. What is particularly important is to align language, education and FDI policies wherever relevant and feasible. For example, English is the official language in Singapore and India, and this language choice has made the two countries an attractive destination for FDI originating from English speaking countries including the UK and US. China has initiated Confucius Institute project to facilitate teaching and learning Chinese language as second language worldwide and Lien, Oh and Selmier (2012) find that such strategy has made positive impact on China's outward FDI. Increasingly, many countries are incorporating English in their curriculum and/or starting to use it as medium of instruction in primary, secondary and tertiary education (Dearden, 2014). Countries in which only vernacular languages are spoken or those use languages different from major FDI source countries could be at a disadvantage in attracting FDI, even if they have a strong HC base. To mitigate the constraint of capitalizing on HC for FDI, policies should be in place to improve linguistic competency of human resources in using languages from major FDI source countries or the most widely spoken lingua franca in international business.

Our findings would also inform senior managers of MNEs that it is insufficient to consider HC and language issues in isolation when making FDI location decisions. Our results suggest that MNEs can leverage HCNs' HC when there is linguistic ease. This has important implications for MNEs investment decisions as well as for subsequent staffing practices in foreign subsidiaries. For example, MNEs that face difficulties in utilizing HCNs may have to opt for home country nationals and third

country nationals, which can be quite an expensive option compared to employing HCNs (Harzing, 2001). Recognizing the importance of language differences in the distribution of international activities around the globe in conjunction with a host country's HC would provide MNEs with locational advantages. Additionally, MNEs should improve linguistic competency within the firms. Those with linguistic capabilities and competent human resources can afford to be more flexible.

Despite its merits, the study has a few limitations. First, we look at FDI as an aggregate construct. However, strategic need for HC and LC in FDI decisions may vary by entry modes, firm strategy and industry. It is possible that wholly owned subsidiaries (WOS) may have different HC and LC issues from international joint ventures (IJVs) as the level of staff interactions and investment in cultural and language training are likely to be higher in WOS than in IJVs (Buckley, Carter, Clegg & Tan, 2005). Equally, Greenfield FDI may have different requirements of HC and LC from mergers and acquisitions (M&As) because Greenfield FDI entails the establishment of new facilities, new routines and new operations, while M&As involve partial or full takeover or the merging with local company, thus there is a need of integrating the acquirer and the acquired firm and their routines and operations. Additionally, due to distinctive characteristics of services, the interactions with the customers are likely to be greater for service firms than for manufacturing firms. As a result, service FDI might have higher level of strategic need for HC and LC than low-tech manufacturing FDI. Future research therefore needs to compare and contrast the integrated effects of HC and LC on shaping decisions of FDI of different types and in different industries.

Second, we operationalize human capital in the form of education and training. However such measures focus on human capital as a single-, macro-level construct (Ployhart & Moliterno, 2011). But human capital can be produced outside formal teaching and training environment provided by schools, universities and firms. Indeed human capital accumulation can vary as a result of family

influence on the knowledge, skills, values and habits (Becker, 1993). Human capital also includes both cognitive and non-cognitive resources. Our measures can only capture cognitive traits, not non-cognitive ones. Future research therefore should apply wider theoretical lenses and multilevel approaches to better capture human capital (Ployhart & Moliterno, 2011). This will allow for investigating the nuances related to the role of human capital in FDI.

The third limitation is related to dataset coverage. Though a large number of countries including developed and developing countries are included in the dataset, our sample does not cover bilateral flows between non-OECD countries. Therefore, the findings may have limitations in explaining the dynamics of FDI flows between developing countries. Fourthly, language variables do not vary over time, which is a common limitation of language statistics. However, some of the language variables may change over time, for example, migration could affect the proportion of population speaking home/host country languages. However, these changes occur slowly. Given our panel covers 14-year time-period, the use of static language variables is reasonable in the context of current study. A related issue is that we focus only on English as lingua franca. Other languages such as French, Spanish and Portuguese are also worthy of examination given their popularity in global commerce. Finally, the measurement of language overlap and religious distance is based on a categorization into groups. Information could be lost from such categorization.

Notes:

[1] FDI is "cross-border investment by a resident entity in one economy with the objective of obtaining a lasting interest in an enterprise resident in another economy" (OECD, 2016).
[2] We recognize that HC can also be acquired through on-the-job training and work experience. However, this is more individual- and firm-specific and is difficult to quantify at the country level.
[3] The firm-level definition of LC is the aggregate possession of relevant language resources of an MNE's employees.
[4] We thank a reviewer for the consideration of composite indices.
[5] Language family affiliations are constructed based on language trees that trace the evolution of languages.
[6] Flow data capture the net value of the capital transactions, in the form of equity investment, reinvested earning and other capital transactions such as intra-company loans, that take place in a given year. Stock data capture the accumulation of flows over time.

[7] We include host and home country dummies when common language, language overlap and linguistic proximity are used. However, host country dummies are not included when English as official language, English as medium of instruction and composite indices are used as English variables are host-country specific variables, correlating with host country dummies.

[8] When the coefficients of *Secondary education* are statistically insignificant, these values are treated as zero when calculating marginal effects.

[9] R^2s for specifications in table 2 increased from 0.06-0.12 to 0.12-0.21 when only positive FDI flows were used. R^2 values further increased to 0.61-0.67 when the models were estimated in logs and are comparable to those studies mentioned above. In terms of the interaction term, we get similar results. However, it is significant only in the specifications with English variables. But as we illustrated before, considering only the positive FDI flows will produce biased results.

[10] Due to space constraint, the results are not presented but available upon request.

Appendix I: List of countries included in the study

Host Countries (81)		Home Countries (88)	
Argentina	Luxembourg	Argentina	Lithuania
Australia	Malaysia	Australia	Luxembourg
Austria	Malta	Austria	Malaysia
Bangladesh	Mexico	Bangladesh	Malta
Belgium	Morocco	Belgium	Mexico
Brazil	Mozambique	Brazil	Morocco
Canada	Nepal	Canada	Mozambique
Chile	New Zealand	Chile	Nepal
China	Nigeria	China	Netherlands
Colombia	Norway	Colombia	New Zealand
Costa Rica	Pakistan	Costa Rica	Nigeria
Croatia	Panama	Croatia	Norway
Czech Republic	Peru	Czech Republic	Pakistan
Denmark	Poland	Denmark	Panama
Ecuador	Portugal	Ecuador	Peru
Egypt	Romania	Egypt	Philippines
El Salvador	Russia	El Salvador	Poland
Estonia	Saudi Arabia	Estonia	Portugal
Ethiopia	Slovakia	Ethiopia	Romania
Fiji	Slovenia	Fiji	Russia
Finland	South Africa	Finland	Saudi Arabia
France	Spain	France	Serbia
Ghana	Sri Lanka	Ghana	Sierra Leone
Greece	Suriname	Greece	Singapore
Guatemala	Sweden	Guatemala	Slovakia
Hong Kong	Switzerland	Hong Kong	Slovenia
Hungary	Syrian Arab Republic	Hungary	South Africa
Iceland	Tanzania	Iceland	Spain
India	Thailand	India	Sri Lanka
Indonesia	Trinidad And Tobago	Indonesia	Suriname
Iran	Turkey	Iran	Sweden
Ireland	United Arab Emirates	Iraq	Switzerland
Israel	UK	Ireland	Syrian Arab Republic
Italy	USA	Israel	Tanzania
Jamaica	Uruguay	Italy	Thailand
Japan	Venezuela	Jamaica	Trinidad And Tobago
Jordan	Vietnam	Japan	Turkey
Kenya		Jordan	United Arab Emirates
Republic of Korea		Kenya	UK
Kuwait		Republic of Korea	USA
Latvia		Kuwait	Uruguay
Lebanon		Latvia	Venezuela
Libya		Lebanon	Vietnam
Lithuania		Libyan Arab Jamahiriya	Zambia

Appendix II: Variable measurement and data sources

Variable	Description/Measurement	Data Source
FDI	Bilateral FDI flow from the home to the host country	OECD's International Direct Investment Database
Common language	A binary variable which takes the value of one if the host and home country share a common language and zero otherwise.	Rose and Spiegel (2011)
Language overlap	Five-point scale based on the incidence (p) of the home country's dominant language(s) in the host country (i.e. the proportion of population (p) in the host country who are able to speak the major language(s) of the home country). 1=p<1%; 2=1% <p<5%; 3=5%<p< 50%; 4=50%<p< 90%; 5= p>90%	Dow and Karunaratna (2006)
English as official language	A binary variable which takes the value of one if English is recognized as an official language in the host country and zero if not.	Central Intelligence Agency (CIA)'s World Factbook
English as medium of instruction	A binary variable which takes the value of one if English is used as a medium of instruction in the host country and zero if not.	'English language proficiency requirement exempted countries' list taken from the website of University of Florida (University of Florida, 2014)
Linguistic proximity	Five-point scale based on the linguistic distance between the two closest major languages for each pair of the host and home country. 1= languages are form different families; 2= languages are from the same family but different branches; 3= languages are from the same branch but different at the first sub-branch level; 4= languages are from the same sub-branch at the first level but different at the second level; 5= languages are same	Dow and Karunaratna (2006)
Secondary education	Secondary School enrolment rate of the host country (% gross)	World Development Indicators
Tertiary education	Tertiary education enrolment rate of the host country (% gross)	World Development Indicators

Appendix II continued

Variable	Description/Measurement	Data Source
Higher education and training	Higher education and training of the host country, a multi-indicator measure of human capital based on secondary and tertiary education enrolment rate, quality of the educational system, math and science education, and the management of schools, internet access in schools, local availability of specialized research and training services, and the extent of staff training.	Global Competitiveness Index (http://www.weforum.org/issues/global-competitiveness)
Population	Population in the host country	World Development Indicators
GDP per capita	GDP per capita in the host country	World Development Indicators
GDP HOME	GDP of the home country	UNCTAD database
Trade openness	Trade openness of the host country, represented by trade intensity: ((exports + imports)/GDP)	World Development Indicators
Infrastructure	Mobile cellular subscriptions per 100 people in the host country	World Development Indicators
Geographical distance	Geographical distance between the host and home country	Rose and Spiegel (2011)
Boarder	A binary variable which takes the value of one if the host and home country share a land border and zero otherwise.	Rose and Spiegel (2011)
RTA	A binary variable which takes the value of one if the host and home country have a regional trade agreement at time t	Rose and Spiegel (2011) for 1995-2006; World Trade Organization (2013) for 2007 and 2008
Cultural distance	A composite measure of cultural distance between the host and home country based on four cultural dimensions: power distance, uncertainty avoidance, individualism, and masculinity	Constructed based on four cultural dimensions taken from the Hofstede Centre (http://geert-hofstede.com/countries.html).
Religious distance	Five-point scale based on the incidence (p) of the home country's dominant religion(s) in the host country: 5=p<1%; 4=1%<p<5%; 3=5%<p<50%; 2=50%<p< 90%; 1=p> 90%	Dow and Karunaratna (2006)
Historical links	A binary variable which takes the value of one if the host and home country were ever in a colonial relationship and zero otherwise.	Rose and Spiegel (2011)

Appendix II continued

Variable	Description/Measurement	Data Source
Legal origin	A binary variable which takes the value of one if the host country has a common-law system (originating in English law) and zero if the host country has a civil law system (French, German, socialist and Scandinavian origins).	La Porta, Lopez-de-Silanes, Shleifer and Vishny (1999)
Market liberalization	A composite measure constructed based on the equal weighted average of following categories from Heritage Foundation Economic Freedom Index: (1) business freedom (2) trade freedom (3) property rights (4) investment freedom and (5) financial freedom.	Heritage Foundation Economic Freedom Index (http://www.heritage.org/index/)
Demographic distance	Differences in demographic characteristics between the host and home country. Demographic characteristics under consideration include: life expectancy at birth; birth rate; the share of population under 14; the share of population above 65	(Berry et al., 2010)
Economic distance	Differences in economic development and macroeconomic characteristics between the host and home country. Economic development and macroeconomic characteristics under consideration include: GDP per capita; exchange rate; export share in GDP; import share in GDP	(Berry et al., 2010)
Financial distance	Differences in financial sector development between the host and home country. Financial sector development variables under consideration include: private sector; stock market size/value; number of listed companies	(Berry et al., 2010)
Knowledge distance	Differences in patents and scientific production between the host and home country. Patents and scientific production variables under consideration include: number of patents per 1 million population; number of scientific articles per 1 million population	(Berry et al., 2010)
Political distance	Differences in political stability, democracy, and trade bloc membership between the host and home country.	(Berry et al., 2010)
Region	Regional dummies. Countries were classified into following regions: Americas, Asia, Europe, Middle-east, Oceania and Africa. Africa is the base region.	Central Intelligence Agency (CIA)'s World Factbook

References

Adserà, A., & Pytliková, M. (2015). The role of language in shaping international migration. *Economic Journal, 125*, F49-F81.

Aichhorn, N., & Puck, J. (2017a). Bridging the language gap in multinational companies: Language strategies and the notion of company-speak. *Journal of World Business, 52*, 386-403.

Aichhorn, N., & Puck, J. (2017b). "I just don't feel comfortable speaking English": Foreign language anxiety as a catalyst for spoken-language barriers in MNCs. *International Business Review, 26*, 749-763.

Anderson, J., & van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review, 93*, 170-192.

Barner-Rasmussen, W., & Aarnio, C. (2011). Shifting the faultlines of language: A quantitative functional-level exploration of language use in MNC subsidiaries. *Journal of World Business, 46*, 288-295.

Barner-Rasmussen, W., & Björkman, I. (2005). Surmounting interunit barriers. *International Studies of Management & Organization, 35*, 28-46.

Becker, G. (1993). *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. Chicago: The University of Chicago Press.

Berry, H., Guillén, M., & Zhou, N. (2010). An institutional approach to cross-national distance. *Journal of International Business Studies, 41*, 1460-1480.

Beugelsdijk, S., Hennart, J.-F., Slangen, A., & Smeets, R. (2010). Why and how FDI stocks are a biased measure of MNE affiliate activity. *Journal of International Business Studies, 41*, 1444-1459.

Blonigen, B. (2005). A review of the empirical literature on FDI determinants. *Atlantic Economic Journal, 33*, 383-403.

Blonigen, B., & Piger, J. (2014). Determinants of foreign direct investment. *Canadian Journal of Economics/Revue canadienne d'économique, 47*, 775-812.

Brannen, M. Y., Piekkari, R., & Tietze, S. (2014). The multifaceted role of language in international business: Unpacking the forms, functions and features of a critical challenge to MNC theory and performance. *Journal of International Business Studies, 45*, 495-507.

Buckley, P. J., Carter, M. J., Clegg, J., & Tan, H. U. I. (2005). Language and social knowledge in foreign-knowledge transfer to China. *International Studies of Management & Organization, 35*, 47-65.

Carstensen, K., & Toubal, F. (2004). Foreign direct investment in Central and Eastern European countries: a dynamic panel analysis. *Journal of Comparative Economics, 32*, 3-22.

Chakrabarti, A. (2001). The determinants of foreign direct investment: sensitivity analyses of cross-country regressions. *Kyklos, 54*, 89-113.

Cheng, L., & Kwan, Y. (2000). What are the determinants of the location of foreign direct investment? The Chinese experience. *Journal of International Economics, 51*, 379-400.

Chiswick, B., & Miller, P. (2005). Linguistic distance: A quantitative measure of the distance between English and other languages. *Journal of Multilingual and Multicultural Development, 26*, 1-11.

Cleeve, E. A., Debrah, Y., & Yiheyis, Z. (2015). Human capital and FDI inflow: An assessment of the African case. *World Development, 74*, 1-14.

Colakoglu, S., Yamao, S., & Lepak, D. P. (2014). Knowledge creation capability in MNC subsidiaries: Examining the roles of global and local knowledge inflows and subsidiary knowledge stocks. *International Business Review, 23*, 91-101.

Contractor, F. J., & Mudambi, S. M. (2008). The influence of human capital investment on the exports of services and goods: An analysis of the top 25 services outsourcing countries. *Management International Review, 48*, 433-445.
Cook, T. R., & Liu, A. H. (2016). Using linguistic networks to explain strength of intellectual property rights. *World Development, 87*, 128-138.
Dearden, J. (2014). *English as a Medium of Instruction - A Growing Global Phenomenon: Phase I Interim Report*. British Council: University of Oxford.
Dhanaraj, C., & Beamish, P. (2009). Institutional environment and subsidiary survival. *Management International Review, 49*, 291-312.
Dow, D., & Karunaratna, A. (2006). Developing a multidimensional instrument to measure psychic distance stimuli. *Journal of International Business Studies, 37*, 578–602.
Dunning, J., & Lundan, S. (2008). *Multinational Enterprises and the Global Economy* (2nd ed.). Cheltenham, UK; Northampton, MA: Edward Elgar.
Eicher, T., Helfman, L., & Lenkoski, A. (2012). Robust FDI determinants: Bayesian Model Averaging in the presence of selection bias. *Journal of Macroeconomics, 34*, 637-651.
Feely, A. J., & Harzing, A.-W. (2003). Language management in multinational companies. *Cross Cultural Management: An International Journal, 10*, 37-52.
Forstenlechner, I., & Mellahi, K. (2011). Gaining legitimacy through hiring local workforce at a premium: The case of MNEs in the United Arab Emirates. *Journal of World Business, 46*, 455-461.
Gao, T. (2005). Labor quality and the location of foreign direct investment: Evidence from China. *China Economic Review, 16*, 274-292.
Globerman, S., & Shapiro, D. (2002). Global foreign direct investment flows: The role of governance infrastructure. *World Development, 30*, 1899-1919.
Gomez-Mejia, L., & Palich, L. (1997). Cultural diversity and the performance of multinational firms. *Journal of International Business Studies, 28*, 309-335.
Gong, Y. (2003). Subsidiary staffing in multinational enterprises: Agency, resources and performance. *Academy of Management Journal, 46*, 728-739.
Grin, F. (1994). The economics of language: Match or mismatch? *International Political Science Review, 15*, 25-42.
Harzing, A.-W. (2001). Who's in charge? An empirical study of executive staffing practices in foreign subsidiaries. *Human Resource Management, 40*, 139-158.
Harzing, A.-W., & Feely, A. (2008). The language barrier and its implications for HQ-subsidiary relationships. *Cross Cultural Management: An International Journal, 15*, 49-61.
Harzing, A.-W., & Pudelko, M. (2013). Language competencies, policies and practices in multinational corporations: A comprehensive review and comparison of Anglophone, Asian, Continental European and Nordic MNCs. *Journal of World Business, 48*, 87-97.
Hejazi, W., & Ma, J. (2011). Gravity, the English language and international business. *Multinational Business Review, 19*, 152-167.
Hitt, M., Biermant, L., Shimizu, K., & Kochhar, R. (2001). Direct and moderating effects of human capital on strategy and performance in professional service firms: A resource-based perspective. *Journal of Management, 44*, 13-28.
Hughes, M., Powell, T. H., Chung, L., & Mellahi, K. (2017). Institutional and Resource-based Explanations for Subsidiary Performance. *British Journal of Management, 28*, 407-424.
Kim, J., & Park, J. (2013). Foreign direct investment and country-specific human capital. *Economic Inquiry, 51*, 198-210.
Kinda, T. (2010). Investment climate and FDI in developing countries: Firm-level evidence. *World Development, 38*, 498-513.

Kostova, T., & Zaheer, S. (1999). Organizational legitimacy under conditions of complexity: The case of the multinational enterprise. *Academy of Management Review, 24*, 64-81.
Ku, H., & Zussman, A. (2010). Lingua franca: The role of English in international trade. *Journal of Economic Behavior & Organization, 75*, 250-260.
La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2008). The economic consequences of legal origins. *Journal of Economic Literature, 46*, 285-332.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization, 15*, 222-279.
La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (2000). Investor protection and corporate governance. *Journal of Financial Economics, 58*, 3-27.
Lewis, M., Simons, G., & Fennig, C. (2014). Ethnologue: Languages of the World. In (17th ed.). Dallas, Texas: SIL International.
Lien, D., Oh, C. H., & Selmier, W. T. (2012). Confucius institute effects on China's trade and FDI: Isn't it delightful when folks afar study Hanyu? *International Review of Economics & Finance, 21*, 147-155.
Luo, Y., & Shenkar, O. (2006). The multinational corporation as a multilingual community: Language and organization in a global context. *Journal of International Business Studies, 37*, 321-339.
Makino, S., & Tsang, E. (2011). Historical ties and foreign direct investment: An exploratory study. *Journal of International Business Studies, 42*, 545-557.
McMahon, W. (2002). *Education and Development: Measuring the Social Benefits*. Oxford: Oxford University Press.
Medvedev, D. (2012). Beyond trade: The impact of preferential trade agreements on FDI inflows. *World Development, 40*, 49-61.
Meier, G. (1995). *Leading Issues in Economic Development*. Oxford: Oxford University Press.
Melitz, J. (2008). Language and foreign trade. *European Economic Review, 52*, 667-699.
Melitz, J., & Toubal, F. (2014). Native language, spoken language, translation and trade. *Journal of International Economics, 93*, 351-363.
Meyer, K. E., Estrin, S., Bhaumik, S. K., & Peng, M. W. (2009). Institutions, resources, and entry strategies in emerging economies. *Strategic Management Journal, 30*, 61-80.
Newburry, W., Gardberg, N. A., & Belkin, L. Y. (2006). Organizational attractiveness is in the eye of the beholder: the interaction of demographic characteristics with foreignness. *Journal of International Business Studies, 37*, 666-686.
Noorbakhsh, F., Paloni, A., & Youssef, A. (2001). Human capital and FDI inflows to developing countries: New empirical evidence. *World Development, 29*, 1593-1610.
Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pacific region. *TESOL Quarterly, 37*, 589-613.
Nunnenkamp, P., & Spatz, J. (2002). Determinants of FDI in developing countries: Has globalization changed the rules of the game. *Transnational Corporations, 11*, 1-34.
Oh, C., Selmier, W., & Lien, D. (2011). International trade, foreign direct investment, and transaction costs in languages. *The Journal of Socio-Economics, 40*, 732-735.
Peltokorpi, V. (2015). Foreign subsidiary top manager nationality and language policy: The moderating effects of subsidiary age and size. *International Business Review, 24*, 739-748.
Peltokorpi, V. (2017). Absorptive capacity in foreign subsidiaries: The effects of language-sensitive recruitment, language training, and interunit knowledge transfer. *International Business Review, 26*, 119-129.
Piekkari, R., Welch, D. E., & Welch, L. S. (2014). *Language in International Business: The Multilingual Reality of Global Business Expansion*. Cheltenham: Edward Elgar.

Ployhart, R., & Moliterno, T. (2011). Emergence of the human capital resource: A multilevel model. *Academy of Management Review, 36*, 127-150.
Reiche, B., Harzing, A.-W., & Pudelko, M. (2015). Why and how does shared language affect subsidiary knowledge inflows? A social identity perspective. *Journal of International Business Studies, 46*, 528-551.
Root, F., & Ahmed, A. (1979). Empirical determinants of manufacturing direct foreign investment in developing countries. *Economic Development and Cultural Change, 27*, 751-767.
Rose, A. K., & Spiegel, M. M. (2011). The Olympic effect. *Economic Journal, 121*, 652-677.
Rugman, A. M. (1986). New theories of the multinational enterprise: An assessment of internalization theory. *Bulletin of Economic Research, 38*, 101-118.
Schneider, F., & Frey, B. (1985). Economic and political determinants of foreign direct investment. *World Development, 13*, 161-175.
Schomaker, M., & Zaheer, S. (2014). The role of language in knowledge transfer to geographically dispersed manufacturing operations. *Journal of International Management, 20*, 55-72.
Schultz, T. W. (1961). Investment in human capital. *American Economic Review, 51*, 1-17.
Selmier, W., & Oh, C. (2013). The power of major trade languages in trade and foreign direct investment. *Review of International Political Economy, 20*, 486-514.
Suliman, A., & Mollick, A. (2009). Human capital development, war and foreign direct investment in Sub-Saharan Africa. *Oxford Development Studies, 37*, 47-61.
Tenzer, H., Pudelko, M., & Harzing, A. (2014). The impact of language barriers on trust formation in multinational teams. *Journal of International Business Studies, 45*, 508-535.
Tenzer, H., Terjesen, S., & Harzing, A.-W. (2017). Language in international business: A review and agenda for future research. *Management International Review, 57*, 815-854.
Uhlenbruck, K. (2004). Developing Acquired Foreign Subsidiaries: The Experience of MNEs in Transition Economies. *Journal of International Business Studies, 35*, 109-123.
Vance, C. M., & Paik, Y. (2005). Forms of host‐country national learning for enhanced MNC absorptive capacity. *Journal of Managerial Psychology, 20*, 590-606.
Vance, C. M., Vaiman, V., & Andersen, T. (2009). The vital liaison role of host country nationals in MNC knowledge management. *Human Resource Management, 48*, 649-659.
Vidal-Suárez, M., & López-Duarte, C. (2013). Language distance and international acquisitions: A transaction cost approach. *International Journal of Cross Cultural Management, 13*, 47-63.
Weitzel, U., Kling, G., & Gerritsen, D. (2014). Testing the fire-sale FDI hypothesis for the European financial crisis. *Journal of International Money and Finance, 49, Part B*, 211-234.
Welch, D., & Welch, L. (2008). The importance of language in international knowledge transfer. *Management International Review, 48*, 339-360.
Welch, D., & Welch, L. (2015). Developing Multilingual Capacity: A Challenge for the Multinational Enterprise. *Journal of Management*.
Wood, G., Mazouz, K., Yin, S., & Cheah, J. (2014). Foreign direct investment from emerging markets to Africa: the HRM context. *Human Resource Management, 53*, 179-201.
Zaheer, S., & Mosakowski, E. (1997). The dynamics of the liability of foreignness: A global study of survival in financial services *Strategic Management Journal, 18*, 439-463.

Table 1: Descriptive statistics and correlation matrix

Variable	Mean	s.d.	Correlation coefficients																				
			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1 FDI	310.78	2484.58																					
2 Secondary education	95.68	22.91	0.02																				
3 Common language	0.12	0.32	0.08	0.06																			
4 Language overlap	1.19	0.63	0.16	0.05	0.58																		
5 English as official language	0.21	0.4	0.1	0.04	0.31	0.31																	
6 English as medium of instruction	0.2	0.4	0.09	0.05	0.31	0.32	0.96																
7 Linguistic proximity	2.05	1.21	0.12	0.07	0.33	0.56	0.26	0.28															
8 Population	63.36	171.68	0.04	-0.29	0.03	0.05	0.18	0.18	-0.02														
9 GDP per capita	16.13	13.14	0.1	0.54	0.05	0.09	0.12	0.09	0.16	-0.14													
10 GDP Home	689.34	1848.91	0.15	-0.09	0.08	0.08	0	0	0.04	0.03	-0.08												
11 Trade openness	86.49	56.29	-0.01	0	-0.06	-0.09	-0.08	-0.17	-0.03	-0.24	0.22	-0.01											
12 Infrastructure	66.99	39.8	0.03	0.41	-0.05	-0.04	-0.06	-0.09	0.03	-0.26	0.49	-0.07	0.39										
13 Religious distance	2.68	1.43	-0.05	-0.11	0.04	-0.09	-0.02	-0.03	-0.36	0.2	-0.1	0.01	-0.03	-0.05									
14 Geographical distance	4130.78	2836.66	-0.05	-0.01	0.16	0.09	0.23	0.23	0.03	0.07	-0.07	0.1	-0.14	-0.09	0.09								
15 Border	0.03	0.18	0.11	0.03	0.04	0.13	-0.06	-0.06	0.09	-0.03	0.02	0	0.01	0.01	-0.11	-0.23							
16 Historical links	0.03	0.18	0.1	-0.02	0.36	0.48	0.11	0.11	0.25	0.01	0.01	0.01	-0.05	-0.03	-0.02	0.02	0.05						
17 RTA	0.52	0.5	0	0.09	-0.04	-0.09	-0.2	-0.18	-0.03	0.01	-0.01	-0.12	0.03	0.09	-0.15	-0.49	0.14	-0.04					
18 Cultural distance	2.92	1.01	-0.08	-0.02	-0.11	-0.17	-0.02	-0.02	-0.09	0.04	0.03	0.02	0.05	0.02	0	0.02	-0.07	-0.11	-0.08				
19 Legal origin	0.24	0.43	0.08	-0.06	0.3	0.28	0.86	0.85	0.22	0.16	0.06	0.03	-0.05	-0.09	0.07	0.24	-0.07	0.1	-0.23	0.02			
20 Market liberalization	68.78	13.73	0.07	0.63	0.12	0.12	0.26	0.22	0.16	-0.33	0.68	-0.09	0.29	0.5	-0.13	0	0.01	0.02	0.04	0	0.15		
21 Tertiary education	52.17	22.93	0.03	0.66	0.1	0.05	0.08	0.09	0.01	-0.27	0.45	-0.09	-0.13	0.51	-0.11	0.05	0.01	-0.03	0.09	-0.02	0	0.54	
22 Higher education and training	4.72	0.89	0.06	0.81	0.08	0.05	0.06	0.06	0.03	-0.14	0.69	-0.11	0.09	0.54	-0.03	-0.06	0.04	-0.01	0.08	0.03	-0.06	0.8	0.75

Note: s.d. indicates standard deviation

Table 2: Baseline estimations with HC measured by *Secondary education* and different LC variables, 1995-2008

LC	Common language	Language overlap	English as official language	English as medium of instruction	English as official language	English as medium of instruction	Linguistic proximity
	(2.1)	(2.2)	(2.3)	(2.4)	(2.5)	(2.6)	(2.7)
Secondary education	5.344	2.894	-2.374*	-2.593**	0.286	0.137	2.478
	(3.270)	(3.253)	(1.278)	(1.323)	(1.194)	(1.224)	(3.146)
LC	-429.4*	-64.22	-448.5***	-329.3**	-373.2***	-310.6**	-61.58
	(241.7)	(145.3)	(134.5)	(133.7)	(132.2)	(132.8)	(57.77)
LCxSecondary education	4.739**	2.549**	8.452***	8.182***	6.829***	6.558***	1.471***
	(2.151)	(1.230)	(2.293)	(2.229)	(1.979)	(1.923)	(0.520)
Population	1.864	1.872	0.493***	0.458***	0.663***	0.648***	1.918
	(1.646)	(1.643)	(0.144)	(0.140)	(0.173)	(0.169)	(1.645)
GDP per capita	80.27***	79.82***	14.81***	14.87***	17.32***	17.46***	79.33***
	(27.59)	(27.62)	(3.864)	(3.843)	(3.815)	(3.808)	(27.53)
GDP Home	0.269**	0.276**	0.116	0.119	0.284**	0.284**	0.270**
	(0.114)	(0.113)	(0.133)	(0.133)	(0.115)	(0.115)	(0.114)
Trade openness	0.809	0.925	-0.669	-0.494	-0.313	-0.187	0.948
	(1.785)	(1.783)	(0.806)	(0.784)	(0.695)	(0.676)	(1.776)
Infrastructure	0.190	0.261	0.525	0.784	-0.781	-0.671	0.169
	(1.023)	(1.015)	(0.897)	(0.905)	(0.932)	(0.941)	(1.022)
Religious distance	17.16	21.22	18.36	21.28	24.93	26.88	30.18
	(20.66)	(21.31)	(19.31)	(19.70)	(20.07)	(20.52)	(21.13)
Geographical distance	-0.0710***	-0.0650***	-0.0583***	-0.0589***	-0.0651***	-0.0658***	-0.0686***
	(0.0140)	(0.0137)	(0.00998)	(0.0100)	(0.0121)	(0.0122)	(0.0140)
Border	1,053***	995.6***	622.9***	626.3***	1,210***	1,211***	1,037***
	(310.1)	(306.5)	(207.2)	(207.3)	(311.6)	(311.5)	(310.4)
Historical links	529.3	269.8	49.67	48.35	629.0	629.0	439.2
	(469.0)	(400.8)	(203.1)	(202.3)	(460.6)	(460.0)	(453.9)
RTA	-188.4*	-201.6*	-103.5*	-113.3**	-121.5*	-127.7*	-189.7*
	(103.3)	(106.1)	(52.89)	(52.93)	(68.14)	(68.91)	(104.3)
Cultural distance	-263.7***	-241.0***	-97.39***	-97.58***	-220.4***	-221.4***	-242.1***
	(81.08)	(74.59)	(20.91)	(20.57)	(54.95)	(54.72)	(79.63)
Legal origin	-712.5	-845.2	163.5***	93.83	206.1***	175.2**	-731.7
	(1,386)	(1,381)	(55.81)	(63.06)	(73.24)	(76.33)	(1,381)
Market liberalization	-5.939	-5.840	2.985	2.863	3.984	3.996	-6.078*
	(3.683)	(3.668)	(2.544)	(2.638)	(2.537)	(2.577)	(3.674)
N	23,987	23,987	18,036	18,036	23,987	23,987	23,987

Country pairs	3,315	3,315	2,430	2,430	3,315	3,315	3,315
R^2	0.122	0.124	0.0643	0.0647	0.106	0.106	0.123
χ^2	348.1***	348.8***	160.8***	161.1***	229.5***	235.2***	345.9***

Notes: Robust standard errors in parentheses. * p<0.1. ** p<0.05. *** p<0.01. Region dummies and country-specific and year-specific fixed effects are not reported for brevity.

Table 3: Marginal effect of HC measured by *Secondary education* under various contexts of LC

LC variable	Value of LC indicator	Context of LC	Marginal effect of human capital (∂ FDI/∂ HC)
Common language	0	Host and home country do not share a common language	0
	1	Host and home country do share a common language	4.739
language proximity	1	p < 1%	2.549
	2	1% ≤ p < 5%	5.098
	3	5% ≤ p < 50%	7.647
	4	50% ≤ p < 90%	10.196
	5	p > 90%	12.745
Linguistic proximity	1	languages are form different families	1.471
	2	languages are from the same family but different branches	2.942
	3	languages are from the same branch but different at the first sub-branch level	4.413
	4	languages are from the same sub-branch at the first level but different at the second level	5.884
	5	languages are same	7.355
Use of English as an official language in the host country	0	English is not used as an official language in the host country	-2.374
	1	English is used as an official language in the host country	6.078
Use of English as a medium of instruction in the host country	0	English is not used as a medium of instruction in the host country	-2.593
	1	English is used as a medium of instruction in the host country	5.589

p = population in host country that can speak the major language(s) in home country

Table 4: Full estimations with HC measured by Higher education and training and different LC variables, 2006-2008

LC	Common language	Language overlap	English as official language	English as medium of instruction	English as official language	English as medium of instruction	Linguistic proximity
	(4.1)	(4.2)	(4.3)	(4.4)	(4.5)	(4.6)	(4.7)
Higher education and training	278.0	-11.14	-149.5	-137.7	-25.84	-8.916	26.27
	(302.0)	(352.5)	(95.36)	(95.07)	(93.46)	(93.46)	(304.2)
LC	-1,493*	-1,144*	-2,063***	-1,874***	-2,015***	-1,759***	-683.6***
	(867.8)	(628.2)	(530.7)	(546.8)	(465.5)	(471.1)	(246.5)
LCxHigher education and training	339.6*	289.1*	434.5***	440.3***	418.9***	405.3***	164.3***
	(195.3)	(148.4)	(131.8)	(137.6)	(114.7)	(119.6)	(57.24)
N	8,557	8,557	6,437	6,437	8,557	8,557	8,557
Country pairs	3,063	3,063	2,269	2,269	3,063	3,063	3,063
R^2	0.146	0.149	0.0732	0.0731	0.130	0.129	0.148
χ^2	297.8***	306.7***	136.4***	136.3***	201.2***	205.3***	300.5***

Notes: Robust standard errors in parentheses. * p<0.1. ** p<0.05. *** p<0.01. Control variables, regional dummies and country-specific and year-specific fixed effects are not reported for brevity.

Table 5: Full estimations with HC measured by tertiary education and different LC variables, 1995-2008

LC	Common language	Language overlap	English as official language	English as medium of instruction	English as official language	English as medium of instruction	Linguistic proximity
	(5.1)	(5.2)	(5.3)	(5.4)	(5.5)	(5.6)	(5.7)
Tertiary education	-7.924***	-10.89*	-3.308*	-3.175*	-1.373	-1.290	-13.96***
	(3.028)	(6.054)	(1.850)	(1.845)	(1.928)	(1.920)	(4.900)
LC	-250.4	7.775	-485.8***	-407.5***	-498.9***	-453.6***	-112.2
	(311.7)	(219.1)	(125.1)	(122.2)	(133.9)	(128.5)	(87.77)
LCxTertiary education	5.546	3.552	19.85***	19.98***	17.63***	17.89***	4.021**
	(5.830)	(4.770)	(4.459)	(4.478)	(3.780)	(3.815)	(1.902)
N	22,722	22,722	17,118	17,118	22,722	22,722	22,722
Country pairs	3,198	3,198	2,342	2,342	3,198	3,198	3,198
R^2	0.122	0.124	0.0704	0.0708	0.109	0.110	0.124
χ^2	343.7***	345.4***	162.6***	162.6***	235.1***	238.6***	343.4***

Notes: Robust standard errors in parentheses. * p<0.1. ** p<0.05. *** p<0.01. Control variables, regional dummies and country-specific and year-specific fixed effects are not reported for brevity.

Table 6: Estimations with composite indices of LC

	LC1	LC2
	(6.1)	(6.2)
Secondary education	1.373	-0.0272
	(1.035)	(1.155)
LC	-125.6*	-338.8***
	(71.60)	(114.2)
LCxSecondary education	3.863***	4.842***
	(0.963)	(1.304)
N	23,987	23,987
Country pairs	3,315	3,315
R^2	0.108	0.106
χ^2	233.9***	231.4***

Notes: Robust standard errors in parentheses. * p<0.1. ** p<0.05. *** p<0.01. Control variables, regional dummies and country-specific and year-specific fixed effects are not reported for brevity.